

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 05-241777

(43)Date of publication of application : 21.09.1993

(51)Int.Cl. G06F 5/00
G06F 15/20
H03M 7/30

(21)Application number : 04-042578

(71)Applicant : FUJITSU LTD

(22)Date of filing : 28.02.1992

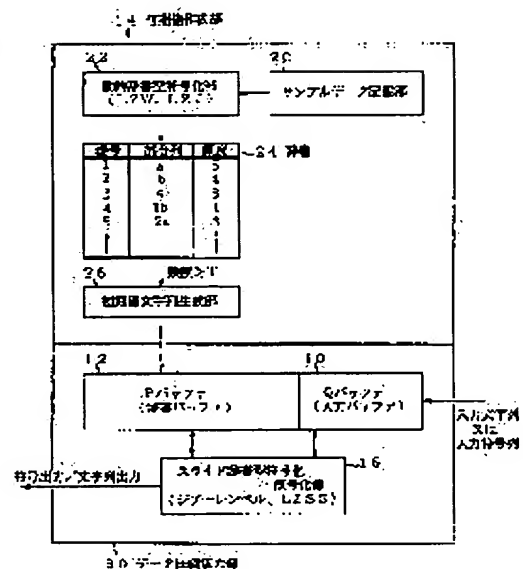
(72)Inventor : YOSHIDA SHIGERU
OKADA YOSHIYUKI
NAKANO YASUHIKO
CHIBA HIROTAKA

(54) DATA COMPRESSION SYSTEM

(57)Abstract:

PURPOSE: To improve the compressibility without spoiling the easiness of slide dictionary type algorithm by registering an initial value character string consisting of a special kind of data which is high in appearance frequency in a dictionary.

CONSTITUTION: An initial value generation part 14 encodes representative sample data according to the dynamic dictionary type algorithm. Then a counter counts the frequency of use of a reference number, indicating a character string registered in a dictionary generated by this encoding, at the time of the encoding as an appearance frequency, and registered character strings in the dictionary 24 which have appearance frequencies larger than a specific threshold value are extracted and arrayed when the encoding of the sample data ends to generate an initialization character string. The initialization value character string generated by the initialization value generation part 14 is used for data compression using a slide dictionary at a data compression and restoration part 30. Thus, the character strings which are high in use frequency are previously registered in the dictionary, so the probability that long input data match a character string to the longest length becomes high to improve the efficiency of the encoding.



LEGAL STATUS

[Date of request for examination] 13.03.1998

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3241788
[Date of registration] 19.10.2001
[Number of appeal against examiner's decision
of rejection]
[Date of requesting appeal against examiner's
decision of rejection]
[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

P-2270

(19)日本国特許庁(JP) (12)公開特許公報(A) (11)特許公開公報番号
特開平5-241777
(43)公開日 平成5年(1993)9月21日

(51)Int.Cl. ⁴	種別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F	5/00	H 9189-5B		
	15/20	G 7343-5L		
H 0 3 M	7/30	8839-5J		

審査請求	未請求	請求項の数 7 (全 23 頁)
(21)出願番号	特願平4-42578	(71)出願人 000005223 富士通株式会社
(22)出願日	平成4年(1992)2月28日	神奈川県川崎市中原区上小田中1015番地 (72)発明者 吉田 茂 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内 (72)発明者 岡田 佳之 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内 (72)発明者 中野 敬彦 神奈川県川崎市中原区上小田中1015番地 富士通株式会社内 (74)代理人 弁護士 竹内 進 (外1名) 最終頁に続く

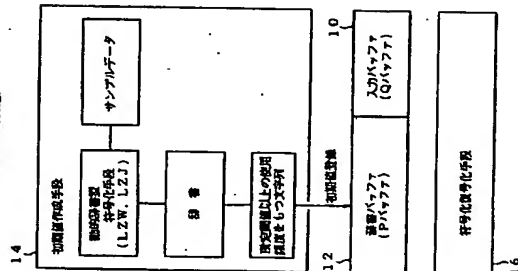
(54)【発明の名称】 データ圧縮方式

(57)【要約】

【目的】 スライド辞書を用いたLZSS符号等によるデータ圧縮方式に関し、スライド辞書型アルゴリズムの効率化を目的とする。

【構成】 代表的なサンプリングデータをLZW符号等の動的辞書型アルゴリズムで符号化して辞書に登録した文字列の使用頻度を求め、所定閾値以上の使用頻度をもつ登録文字列から初期値文字列を作成してスライド辞書に初期登録して符号化を開始する。

本発明の原理図説明



【特許請求の範囲】

【請求項1】 入力バッファ(110)中の入力データを辞書バッファ(112)中の符号化済みデータの部分列の内、最長一致するものの格納位置と一致して指定して符号化し、符号化済みの入力データを辞書バッファ(112)に移して新たな符号化済みデータとして次の入力データを符号化するデータ圧縮方式に於いて、

代表的なサンプリングデータを相異なる部分列に分けたとき、所定閾値以上の出現頻度をもつ部分列を抽出し、該抽出した部分列を一列に並べて初期値文字列を予め作成する初期値作成手段(114)と、

該初期値作成手段(114)で作成した初期値文字列を、符号化又は復号化に先立って最初に前記辞書バッファ

(112)に自動的に設定して符号化又は復号化済みデータと見做し、該初期値文字列および新たに設定した符号化済みデータと最長一致する部分列を探索して

(110)の入力データと最長一致する部分列を探索して格納位置と一致して指定して符号化又は復号化する符号化復号化手段(116)と、を備えたことを特徴とするデータ圧縮方式。

【請求項2】 入力バッファ(110)中の入力データを辞書バッファ(112)中の符号化済みデータの部分列の内、最長一致するものの格納位置と一致して指定して符号化し、符号化済みの入力データを辞書バッファ(112)に移して新たな符号化済みデータとして次の入力データを符号化するデータ圧縮方式に於いて、

代表的なサンプリングデータを相異なる部分列に分けたとき、所定閾値以上の出現頻度をもつ部分列を抽出し、該抽出した部分列を一列に並べて初期値文字列を予め作成する初期値作成手段(114)と、

該初期値作成手段(114)で作成した初期値文字列を、符号化に先立って最初に前記辞書バッファ(112)に自動的に設定して符号化済みデータと見做し、該初期値文字列の中からのみ前記入力バッファ(110)の入力データと最長一致する部分列を探索して格納位置と一致して指定して符号化又は復号化する符号化復号化手段(116)と、

を備えたことを特徴とするデータ圧縮方式。

【請求項3】 請求項1、2記載のデータ圧縮方式に於いて、前記初期値作成手段(114)は符号化済み文字列を参照番号を付して登録する辞書を有し、代表的なサンプリングデータの文字列に最長一致する前記辞書中の符号化済み部分列を探索して参照番号で指定して符号化し、該符号化後に該参照番号に次のサンプリング文字を付加した部分列を新たな参照番号を付して前記辞書に登録し、更に辞書に登録された符号化済み文字列の探索する毎に使用回数を計数し、前記サンプリングデータの符号化終了時に前記使用回数が所定閾値以上となる文字列を抽出し、該抽出した文字列を出現順に並べて初期値文字列を予め作成することを特徴とするデータ圧縮方式。

【請求項4】 請求項1、2記載のデータ圧縮方式に於いて、前記初期値作成手段(114)は符号化済み文字列を参照番号を付して登録する辞書を有し、代表的なサンプリングデータの文字列に最長一致する前記辞書中の符号化済み部分列を探索して参照番号で指定して符号化し、該符号化後に該参照番号に次のサンプリング文字を付加した部分列を新たな参照番号を付して前記辞書に登録し、更に辞書の部分列を探索して符号化済み部分列の探索する毎に使用回数を計数し、前記サンプリングデータの符号化終了時に前記使用回数が所定閾値以上となる部分列を抽出し、該抽出した部分列を出現順に並べて初期値文字列を予め作成することを特徴とするデータ圧縮方式。

【請求項5】 請求項1、2記載のデータ圧縮方式に於いて、前記初期値作成手段(114)は入力バッファ中の代表的なサンプリングデータを辞書バッファ中の符号化済みデータの部分列の内、最長一致するものの格納位置と一致して指定して符号化し、該符号化データを相異なる部分列に分けたとき、所定閾値以上の出現頻度をもつ部分列を抽出し、該抽出した部分列を出現順に並べて初期値文字列を予め作成することを特徴とするデータ圧縮方式。

【請求項6】 請求項1、2記載のデータ圧縮方式に於いて、前記辞書バッファ(112)を抽出専用の第1メモリと、読み及び書き可能な第2メモリで構成し、前記初期値作成手段(114)で作成した初期値文字列を前記第1メモリに固定行に記憶し、前記入力バッファ(110)の符号化済み入力データは第2メモリに移して記憶することを特徴とするデータ圧縮方式。

【請求項7】 請求項1、2記載のデータ圧縮方式に於いて、前記辞書バッファ(112)を辞書登録禁止領域を設定した読み及び書き可能なメモリで構成し、符号化開始時に前記初期値作成手段(114)で作成した初期値文字列を前記辞書バッファ(112)の辞書登録禁止領域にロードした後に符号化を開始することを特徴とするデータ圧縮方式。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、スライド辞書を用いたジブレンベル符号化によるデータ圧縮方式に関し、特にスライド辞書を用いた符号化に動的辞書を用いた符号化を利用したデータ圧縮方式に関する。近年、文字コード、ペグトル情報、画像など様々な種類のデータがコンピュータで扱われるようになっており、扱われるデータも急速に増加してきている。

【0002】 大量のデータを扱うときは、データの中の冗長部分を省いてデータ量を圧縮することで、記憶容量を減らしたり、速く伝送したりできることになる。様々なデータを1つの方式でデータ圧縮できる方法としてユニバーサル符号化が提案されている。ここで、本発明の分野は、文字コードの圧縮に限らず、様々なデータに

適用できるが、以下では、情報理論で用いられている呼称を踏襲し、データの1ワード単位を文字と呼び、データが任意ワードつながったものを文字列と呼ぶことにする。

[00003]ユニバーサル符号の代表的な方法として、ジブレンベル(Ziv-Lempel)符号がある(詳しくは、例えば、森像「Ziv-Lempel」符号があるデータ圧縮法」、情報処理、Vol. 26, No. 1, 1985年を参照のこと)。ジブレンベル符号では①スライド辞書型(ユニバーサル型という)と、②動的辞書型(増分分解型という)の2つのアルゴリズムが提案されている。これらの方式の両方への改良方法が発見され、補助記憶装置のファイル圧縮や、パソコン通信でのデータ伝送に利用されるようになっていく。

[0004] [従来の技術] まず従来のスライド辞書型のアルゴリズムと動的辞書型のアルゴリズムについて説明する。

Pバッファ中の最長一致系列の開始位置

[0007] 次にQバッファ10内の符号化した文字列をPバッファ12に移して、新たなデータを得る。以下、同様の操作を繰り返して、データを部分別に分割して符号化する。すなわちジブレンベル符号では、現在の文字コードの格納を、符号化済の過去の系列からの複製として符号化するものである。ジブレンベル符号を用いた場合、文字コードの文書情報は1/2程度に圧縮できる。

[0008] 更にスライド辞書型アルゴリズムの改良として、LZSS符号がある(T. C. Bell, 'Beller OPM/LT ext-Compression', IEEE Trans. on Commun., Vol. COM-34, No. 12, Dec. 1986参照)。LZSS符号では、[Pバッファ中の最大一致系列の開始位置]と「一致する長さ」の組を、「次のシンボル」とをフラグ区別して、符号量の少ない方で符号化する。

[0009] 更にスライド辞書型アルゴリズムの改良として、LZS符号がある。LZS符号は、LZSS符号の簡易版であるQ1C-122符号がある。LZS符号について次に説明する。

[LZS符号] LZS符号による符号化の処理フローを図11に示し、その原理図を図12及び図13に示す。

[0010] LZS符号による符号化は、図12(b)に示すように例えば4ビットのインデックス情報をもって、これから符号化する文字列を格納する例えば4ビットのインデックス情報に対処して16個の文字数

(1) スライド辞書型アルゴリズム

このアルゴリズムは、演算量が多いが、高圧縮率が得られる方法である。即ち、符号化データを、過去のデータ系列の任意の位置から一致する最大長の系列に区切り(部分列)、過去の文字列の複製として符号化する方法である。

[ジブレンベル符号] 図10にスライド辞書型であるジブレンベル符号の符号器の原理図を示す。

[0005] 図10において、辞書バッファとしてのPバッファ12には符号化済みの入力データが格納されており、入力バッファとしてのQバッファ10にはこれから符号化するデータが入力されている。Qバッファ10の文字列をPバッファ12の文字列と照合し、Pバッファ12の中で一致する最大長の文字部分列を求め、そして、Pバッファ12中でこの最大長文字列を指定する。ため次の情報の組を符号化する。

[0006] [表1]

一致する長さ

不一致のシンボル

を格納できるQバッファ10と、図12(a)に示すように、例えば12ビットのインデックス情報をもって4096個の符号化済の文字列を格納するPバッファ12とを備えるようにして構成する。

[0011] 符号化処理は図11のフローチャートに示すように、ステップS1でPバッファ12を空にしてQバッファ10に入力データを格納した後に、ステップS2でQバッファ10の文字列とPバッファ12の文字列とを照合し最長一致する文字部分列を求め、ステップS3で2文字以上であることを条件にステップS5に進んで、求められた文字部分列を指定するために[文字列Sの出現位置]

[一致長]の組で符号化する。

[0012] 続いてステップS6でQバッファ10内の符号化した文字列をPバッファ12に移して、Qバッファ10内に符号化した文字列分の新たな文字列を入力していくことで符号化を繰り返す。尚、最長一致文字部分列が1バイトのときは生データで符号化した方が有利であるので、ステップS4で

生データ1バイト)をそのまま出力する。

[0013] 更に、図13に示すように、8個の符号化データもしくは生データを1組のデータとしてまとめると共に、まとめられた各8個のデータが符号化データなのか生データなのかを示すステップS4、S5で得られたフラグビットとなる8ビット識別データを先頭に付加し、1組のデータとして出力する。

(2) 動的分解型(増分分解)アルゴリズム

このアルゴリズムは、圧縮率はユニバーサル型より劣るが、シンプルで、計算も容易であることが知られている。

[0014] 増分分解型ジブレンベル符号では、入力シンボルの系列を

$X = aababababaa \dots$

とすると、成分系列 $X = X_0 X_1 X_2 \dots$ への増分分解は次のようになる。まず X_1 を既成分の右端のシンボルを取り除いた最長の列とし、

$X = a \cdot ab \cdot abab \cdot b \cdot a \cdot a \cdot a \cdot \dots$

となる。従って、

$X_0 = a$ (空列)

$X_1 = X_0 b$

$X_2 = X_1 a$

$X_3 = X_2 a$

$X_4 = X_3 b$

$X_5 = X_4 a \dots$

と分解できる。増分分解した各成分系列は既成分系列を用いて次のような組で符号化する。

[0015] [表2]

成分のインデックス
(各成分の出た順番)

次のシンボル

[0016] すなわち、増分分解型アルゴリズムは、符号化パターンについて、過去に分解した部分列の内、最長一致するものを求め、過去に分解した部分列の複製として符号化するものである。動的辞書型アルゴリズムの改良としては、

①LZW (Lempel-Ziv-Weich) 符号 (T. A. Welch, 'A Technique for High-Performance Data Compression', Computer, June 1984 参照)

②LZJ 符号 (J. Jakobsson, 'Compression of Character Strings by An Adaptive Dictionary', BIT, 25th, 1985年, 593-603頁参照のこと)

とがある。次にLZW符号について説明する。

[LZW符号] LZW符号の符号化の処理のフローを図14に示す。即ちLZW符号化は、書き換え可能な辞書をもち、入力文字コードのデータ中を相異なる文字列に分け、入力文字コードのデータ中を相異なる文字列に分け、この文字列が出現した際に番号を付けて辞書に登録すると共に、現在入力している文字列を辞書に登録し、ある最長一致文字列の番号だけで表して、符号化の技ものである。尚、動的辞書型符号およびLZW符号の技術は、特開昭59-231683、米特許4,558,302で開示されている。図14の符号化処理は次のようになる。

[0017] S1: 予め全文字につき一文字からなる文字列を初期値として登録してから符号化を始める。辞書の登録数nを文字数Aと置く。カーソルをデータの先頭の位置に置く。

S2: カーソルの位置からの文字列に一致する辞書登録の最長文字列Sを見つける。

[0018] S3: 文字列Sの辞書番号を $(\log_2 n)$ ビットで表して出力する。ただし、 $(\log_2 n)$ は $\log_2 n$ 以上の最小の整数である。辞書登録数nを一つインクリメントする。

S4: 文字列Sにカーソルの最初の文字Cを付加した文字列SCを辞書に登録する。カーソルは文字列Sの後の文字に移動させる。S2に戻る。

[0019] 図15はLZW符号の符号化を示したフローチャートであり、符号化の逆の処理となる。動的辞書型アルゴリズムは、辞書内の系列は過去に符号化した(サンプリングした)系列の中だけから選ぶため、処理速度が遅い。しかし、過去に現れたデータの一部の系列しか含まないため圧縮率が高く取れない欠点がある。

[0020] 動的辞書型アルゴリズムの改良版として、辞書への学習量を増やしインデックスのみで符号化できるようにしたLZJ符号がある。

[LZJ符号] LZJ符号の符号化の処理フローを図16に示し、また復号化の処理フローを図17に示す。

[0021] ここで、辞書と文字列の表記法を次のように定義する。文字列の集合をAとし、集合Aの文字を組み合わせでできる文字列をSで表す。文字列Sのi番目の文字を $S(i)$ とする。更に複数の部分文字列S

$(i), S(i+1), \dots, S(j)$ を $S(i, j)$ とする。辞書を $D_h(S)$ で表し、辞書の木(iree)の根(root)から葉(leaf)へのパスとして文字列S中の一定の長さhの全ての部分文字列を登録する。

[0022] 図16のLZJ符号化処理は次のようになる。

S1: 辞書に全文字の一文字を初期値として登録してから符号化を始める。辞書の登録数nを文字数Aとおく。カーソルk=0とおく。

S2~S5: k番目の入力文字まで符号化が終了したとして文字列 $S(i, k)$ の全ての部分文字列がすでに辞書 $D_h(S(i, k))$ に登録してある。S(k+1), ..., の文字列から符号化する。

[0023] 詳細に説明すると、次のようになる。
S2: $S(k+1), \dots$ から辞書 $D_h(S(i, k+1))$ の登録文字列に最長一致する部分文字列 $S(k+1, k+2)$ を見つける。

S3: 部分文字列 $S(k+1, k+2)$ の辞書番号 a_x を $(\log_2 n)$ ビットで表して出力する。ただし、nは辞書の現在の登録数であり、 $(\log_2 n)$ は $\log_2 n$ 以上の最小の整数である。ここで、符号番号 a_x は部

分文字列 $S(i_x, j_x)$ を設す。各々の a_x は辞書D $h(S(i_x, j_x))$ 、 $(i_x \leq j_x \leq i_x + h, i_x = j_x - 1 + 1)$ の辞書番号である。
[0024] S4: 部分文字列 $S(k-h+2, k+1), \dots, S(k+h-1, k+z)$ に n をインクリメントしながら辞書番号を付けて辞書に追加し、辞書D $h(S(i, k+z))$ を構成する。
S5: カーソル $k=k+z$ とおく。
S6: 全文字列を処理するまでS1~S5を繰り返す。
[0025] ここでステップS4の文字列の辞書登録を図示すると図18に示すようになる。次に図17のLZJ復号処理は次のようになる。
S1: 図16のS1と同様に辞書に全文字列の一字を初期値として登録する。辞書の登録数 n を文字種数 A とおく。カーソル $k=0$ とおく。
[0026] S2~S4: 辞書番号 a_w が復号化され、文字列 $S(i, j_w)$ まで利用することができ、辞書D $h(S(i, j_w))$ が再構成されている。次に辞書 a_{w+1} を復号する。詳細に説明する次のようになる。
S2: 辞書 a_{w+1} を復号した辞書番号より辞書D $h(S(i, j_w))$ 内の部分列 $S(i_{w+1}, j_{w+1})$ を復元する。部分列 $S(i_{w+1}, j_{w+1})$ は辞書内で根 $(root)$ からアドレス a_{w+1} の節点で取られる文字列である。
[0027] S3: 文字列 $S(i, j_{w+1})$ を復号した後、辞書D $h(S(i, j_{w+1}))$ を図16のS4と同様に構成する。
S4: カーソル $k=j_{w+1}$ とおく。
S5: 全符号を処理するまでS1~S4を繰り返す。
[0028]

[発明が解決しようとする課題] しかしながら、従来のスライド辞書型アルゴリズムのLZS符号および動的辞書型アルゴリズムのLZW符号は完全なユニバーサル性を前提にしており、辞書が空白の状態から符号化されるようにしている。このため、従来の符号化方式では、入力データの始めの方で、学習量が少ない(辞書内容が少ない)とき、圧縮率が低いという欠点があった。
[0029] LZW符号ではユニバーサル性も重要であるが、入力データに特定の種類のデータだけに多く現れるときは、辞書は必ずしも空白の状態から符号化される必要はない。この観点から本発明者は、動的辞書型アルゴリズムにおいて図19に示すように、高い頻度で出現する文字列のみ保持した辞書を用いて高圧縮率を得る方法を提案している。

[0030] 図19においては、サンプリングデータを対象にLZW符号化を行って辞書を作成し、辞書には使用頻度を同時に計数しておく。サンプリングデータの符号化が済んだ場合には、辞書の中から出現頻度が閾値T以上の文字列を抽出して実際の符号化に使用する辞書に初期値として登録してから符号化又は復号化を行う。しかし動的

辞書型アルゴリズムは符号化、復号化の処理速度はバランしている。一方、スライド辞書型アルゴリズムにおいては、符号化は速いものの復号化の処理が格段に遅いため、データベースなどのように復元の処理が主になる用途においては不利となる。

[0031] 本発明は、このような状況に鑑みてなされたもので、動的辞書型アルゴリズムを利用した初期値の登録でスライド辞書型アルゴリズムによる符号化を効率よくできるようにしたデータ圧縮方式を提供することを目的とする。

[0032] [課題を解決するための手段] 図1は本発明の原理説明図である。まず本発明は、入力バッファ(Qバッファ)12の入力データデータを辞書バッファ(Pバッファ)12中の符号化済データの部分列の内、最長一致するものの格納位置と一致長で指定して符号化し、符号化済みの入力データは辞書バッファ12に移して新たな符号化済みデータとして次の入力データを符号化するスライド辞書型アルゴリズムに従ってデータ圧縮方式を対象とする。

[0033] このようなデータ圧縮方式につき本発明は、あつては、代表的なサンプリングデータを最長の部分列に分けたとき、所定閾値T以上の出現頻度をもつ部分列を抽出し、この抽出した部分列を1列に並べて初期値文字列を予め作成する初期値作成手段14と、初期値作成手段14で作成した初期値文字列を、符号化又は復号化に先立って最初に辞書バッファ12に自動的に設定して符号化又は復号化済データと見做し、初期値文字列および新たな符号化済データと見做し、初期値文字列から入力バッファ12の入力データと最長一致する部分列を検索して格納位置と一致長で指定して符号化又は復号化する符号化復号化手段16とを備えたことを特徴とする。

[0034] また本発明は、初期値作成手段14で作成した初期値文字列を、符号化に先立って最初に辞書バッファ12に自動的に設定して符号化済又は復号化済データと見做し、初期値文字列の中からのみ入力バッファ12の入力データと最長一致する部分列を検索して格納位置と一致長で指定して符号化し、辞書バッファ12に対する新たな入力データの登録は行わないようにしてもよい。

[0035] ここで、初期値作成手段14は動的辞書型アルゴリズムであるLZW符号の符号化処理に従って初期値文字列を作成する。即ち、符号化済み文字列を参照番号を付して登録する辞書を有し、代表的なサンプリングデータの文字列に最長一致する辞書中の符号化済み部分列を検索して参照番号で指定して符号化し、この符号化後に参照番号に次のサンプリング文字を付加した部分列を新たな参照番号を付して辞書に登録し、更に辞書に登録された符号化済み文字列の検索する際に使用頻度を計数し、前記サンプリングデータの符号化終了した際に前記使用頻度

が所定閾値以上となる文字列を抽出し、抽出した文字列を出現順に並べて初期値文字列を予め作成する。
[0036] また初期値作成手段14は動的辞書型アルゴリズムであるLZW符号の符号化処理に従って初期値文字列を作成してもよい。即ち、初期値作成手段14は符号化済み文字列を参照番号を付して登録する辞書を有し、代表的なサンプリングデータの文字列に最長一致する辞書中の符号化済み部分列を検索して参照番号で指定して符号化し、この符号化後に符号化した入力文字列の各文字を順次後部部分列とし、この後部部分列に辞書中の部分列を加えて一定長の部分列を複数作成して全て辞書に登録し、更に辞書に登録された符号化済み部分列の検索する際に使用頻度を計数し、サンプリングデータの符号化終了した際に使用頻度が所定閾値以上となる部分列を抽出し、この抽出した部分列を出現順に並べて初期値文字列を予め作成する。

[0037] 更に初期値作成手段(14)は、サンプリングデータをスライド辞書型アルゴリズムで符号化して符号列を求め、この符号列に動的辞書型アルゴリズムであるLAW符号化しLZW符号化して符号化して使用頻度の高い初期値文字列を生成してもよい。即ち、入力バッファ中の代表的なサンプリングデータの文字列を辞書バッファ中の符号化済データの部分列の内、最長一致するものの格納位置と一致長で指定して符号化し、この符号化データを相異なる部分列に分けたとき、所定閾値以上の出現頻度をもつ部分列を抽出し、抽出した部分列を出現順に並べて初期値文字列を予め作成する。

[0038] 学習により作成した使用頻度の高い初期値文字列の辞書バッファ12への登録は、辞書バッファ12を読出専用の第1メモリ(ROM)と、読出及び書込可能な第2メモリ(RAM)で構成し、初期値作成手段14で作成した初期値文字列を第1メモリ(ROM)に自動的に記憶し、入力バッファ12の符号化済み入力データは第2メモリ(RAM)に移して記憶する。
[0039] また辞書バッファ12を学習禁止領域を設定した読出及び書込可能なメモリ(RAM)で構成し、符号化開始時に初期値作成手段14で作成した初期値文字列を辞書バッファ12の学習禁止領域にロードした後に符号化を開始するようにしてもよい。

[0040] [作用] このような構成を備えた本発明のデータ圧縮方式によれば、次の作用が得られる。まず圧縮符号化しようとするデータの種類の数に同じサンプリングデータを対象に動的辞書型アルゴリズム、即ちLZW符号やLZW符号のアルゴリズムに従った符号化を行うと共に、この符号化に使用する辞書にカウンタを設け、参照番号が符号化に使われた回数を使用頻度として計数する。

[0041] サンプリングデータの符号化が済んだならば、辞書の登録済み文字列の気候風の接点に設けた使用頻度を示すカウンタ計数値の小さい文字列は辞書から削除

し、高頻度で出現する文字列のみを残した辞書求め、この文字列から一列に並べて初期値文字列を生成する。学習により生成された初期値文字列を用いたスライド辞書型アルゴリズムによる符号化は、次のようにする。
[0042] ①符号化時に、予め記憶装置に取り出しておいた高頻度の初期値文字列を初期値として辞書にロードした後、符号化する。

②予め作成した高頻度の文字列を初期値として辞書の先頭に書き替えるように固定部分として設定しておき、符号化する。スライド辞書型アルゴリズム、例えばLZS符号による符号化であっても、使用頻度の高い文字列S符号による符号化に初期値として登録されているが、最初から長い入力データの文字列に最長一致する辞書が予め辞書バッファに初期値として登録されているため、最初から長い入力データの文字列に最長一致する辞書の文字列を登録することができ、スライド辞書型の符号化をより高速で行うことができる。

[0043] [実施例] 図2は本発明の一実施例を示した流れ図構成図である。図2において、14は初期値作成部であり、代表的なサンプリングデータを対象に動的辞書型アルゴリズムに従った符号化を行い、この符号化で作成される辞書に登録された文字列(部分列)を示す参照番号の符号化時に使われた回数を出現頻度としてカウンタで計数し、サンプリングデータの符号化が終了した時点で所定閾値以上の出現頻度をもつ辞書24の登録文字列(部分列)を取り出し、一列に並べて初期値文字列を作成する。

[0044] 具体的には、初期値作成部14はサンプリングデータ記憶部20、動的辞書型符号化部22、辞書24及び初期値文字列生成部26で構成される。サンプリングデータ記憶部20にはデータ圧縮の対象となるデータの種類の数に同じ代表的なサンプリングデータが学習対象として記憶される。動的辞書型符号化部22は動的辞書型アルゴリズムに従ってサンプリングデータ記憶部20のサンプリングデータを対象に辞書24を作成しながら符号化を行う。

[0045] この動的辞書型アルゴリズムとしては、例えば図13に示したLZW符号化アルゴリズム、あるいは図15に示したLZW符号化アルゴリズムを用いることができる。辞書24は参照番号に対応して符号化済み文字列としての部分列を登録しており、更に符号化済み部分列の参照番号が符号化時に使われた回数を計数するカウンタを設け、出現頻度として計数するようにしている。初期値文字列生成部26は動的辞書型符号化部22によるサンプリングデータの符号化処理が終了した段階で辞書24の中から所定の閾値T、例えばT=2以上となる出現頻度をもつ部分列を取り出し、この部分列を一列に並べて使用頻度の高い初期値文字列を生成する。

[0046] 初期値作成部14で作成される初期値文字列はデータ圧縮部30におけるスライド辞書型アルゴリズムを用いられる。データ圧縮部30は入力バッファとしてのQバッファ10、辞書バッファとしてのPバッファ12及びスライド辞書型符号化復号化部

1.6で構成される。Pバッファ1.2には符号化及び復号化に先立って初期値作成部1.4で予め作成された初期値文字列が登録され、この初期値の登録領域はQバッファ1.0から入力文字列をシフトして格納しても変更されず、恒定的に保持される。即ち、Pバッファ1.2に登録した初期値文字列を符号化済みデータと見做してスライド辞書型アルゴリズムに従った符号化及び復号化を行う。

【0047】スライド辞書型符号化復号部1.6はスライド辞書型アルゴリズムに従った符号化または復号化を行うもので、具体的には、ジブレンベル符号化アルゴリズムや図1.0に示したLZSS符号化アルゴリズムを実行する。図3は図2の初期値作成部1.4の処理を示したフローチャートである。図3における初期値作成処理は次のようになる。

【0048】S1:入力データとして多く出現するデータサンプルを入力し、LZW符号またはLZJ符号等の動的辞書型アルゴリズムに従ってサンプルデータの符号化を行う。この符号化において、木構造の辞書が作成される。同時に符号化により作成される辞書の木構造における各節点が文字列を収容することになるが、各節点にカウントを付加して符号化時に最長一致する文字列を検出し、たときに通過した各節点でカウントを1つインクリメントして使用回数を計数する。

【0049】即ち、最長一致の文字列を検出した場合には、検索した文字列に含まれる節点のカウントの全てがカウントアップされることになる。

S2:サンプルデータの符号化が終了した時点で辞書の各節点の連鎖で構成される文字列の中から節点に設けたカウントの計数値が所定の閾値T以上の高頻度で使われた文字列を取り出す。

【0050】S3:S2で取り出した文字列を1列の文字列の形に並べ、初期値文字列を生成する。このとき既に並べた文字列の中に新たに取り出した文字列と同じ文字列があるか否かを検索し、もし同じ文字列があれば重複するので、初期値の文字列には含めないようにする。以上のS1～S3の処理を経て作成された初期値文字列はスライド辞書型アルゴリズムに従った符号化及び復号化に使用するため、外部の補助記憶装置等に取り出しておくことが望ましい。

【0051】図4は図3の初期値処理のステップS1における動的辞書型アルゴリズムに従ったサンプルデータの符号化の際に作成される辞書の木構造と各節点に設けたカウントによる使用頻度の計数を示した説明図である。図4には、まず辞書に例えばa b c dの各文字を参照番号①～④に示すように初期登録した後にサンプルデータの符号化を開始しており、図示の状態においては更に参照番号⑤～⑨までの文字列の登録が済んだ状態で木構造を示している。

【0052】例えば、入力データa b cの符号化は参照

番号①～⑦で示す文字列の登録が済んだ段階で行われており、文字列a bは辞書2.4の検索により参照番号⑥で示す文字列に一致していることから、出力符号は「⑥c」として出力し、続いて参照番号⑥に次の1文字cを加えた文字列を新たに参照番号⑧を付して辞書2.4に登録している。

【0053】次の文字列a b dについても、辞書2.4の検索で参照番号⑥の文字列に最長一致し、従って出力符号は「⑥d」として出力され、辞書2.4に対しては参照番号⑥に次の1文字dを加えた文字列に新たな参照番号⑩を加えて登録している。この文字列a b cと文字列a b dの符号化後の登録にあっては、同じ文字列a bが2回使用されているため、参照番号⑩の節点及び参照番号⑩の節点の各カウントが2回カウントアップされ、それぞれ5、3となる。

【0054】このような木構造を備えた辞書2.4における各節点の計数値は、その節点の子供の計数値の和に1を加えた値となる。例えば、参照番号⑩の文字aの節点の計数値はその節点の子供となす参照番号⑩と⑥の計数値の和1+3=4に1を加えた値として5となる。図5は図3のステップS2及びS3に示したサンプルデータの符号化で得られた辞書から閾値T以上の高頻度で使われた文字列を取り出して、初期値として使用する文字列の形に変換する処理を示した説明図である。

【0055】図5(a)はサンプルデータの符号化が終了した状態で得られた辞書の木構造を示したもので、各文字の節点に設けた辞書内の計数値が使用頻度を示している。この図5(a)の辞書の木構造に対し、閾値T=2以上の計数値をもつ文字列を取り出すと、図5(b)に示すようになる。この閾値2以上の計数値をもつ文字列は、図5(c)に示すように、例えば左側の文字列から順番に1本の文字列にまとめるように並べ替えられ、スライド辞書型アルゴリズムのPバッファに初期値として登録される初期値文字列が作られる。

【0056】図6は図2のデータ圧縮部3.0におけるスライド辞書を用いた符号化処理を示したフローチャートであり、次のようになる。

S1:図3の初期値作成処理で作られた初期値文字列(N文字)をPバッファの前部に格納する。

S2:通常のスライド辞書型アルゴリズムと同様に、入力したQバッファ1.0の文字列を初期値文字列と符号化済み文字列を格納したPバッファ1.2から最長一致する文字列を検索して、開始位置と一致長の組で符号化する。この場合、符号化文字列は2文字一致するか否かによって次の2つのモードで符号化される。

【0057】符号化モード【識別ビット0】【最長一致文字列の位置】【一致長】

生データモード【識別ビット1】【1文字】

S3:辞書の削除及び登録処理として初期値文字列を除くPバッファ1.0内の部分をスライドさせる。即ち、P

バッファ1.0内の初期値文字列を示す0～n-1の位置の文字はそのままだにしておき、文字位置nから符号化が済んだQバッファ1.0の文字数だけ文字列を左にシフトし、Pバッファ1.2の左側から右側にシフトして追加する。

【0058】以上の辞書削除及び登録処理が済んだならばQバッファ1.0に符号化済みの文字列を左にシフトし、新たな文字列を入力する。以下同様にして、ステップS2、S3の処理を繰り返して入力文字列を符号化する。図7は図6のSライド辞書型アルゴリズムを用いた符号化におけるPバッファの構成を示した説明図である。

【0059】図7(a)はPバッファ1.2として書き換え可能なメモリ、例えばRAMを使用した場合であり、Pバッファ1.2の前部の領域で示す0～n-1の領域に初期値文字列をロードし、残りを符号化済み文字列の登録領域としておく。この初期値文字列ロード領域0～n-1についてはその後の書き換えを禁止し、新たなPバッファ1.2に対する登録に際しては、nの位置から、右から左に登録済み文字列のシフトに伴って、登録していた文字列を逐次左にシフトして格納して削除する。

【0060】図7(b)はPバッファ1.2の他の構成を示したもので、この実施例においては、初期値文字列を登録する領域として固定記憶を実現するため、ROMを用いたROM領域としており、残りの領域を書き換え可能なRAM領域を用いたRAM領域を設けている。Pバッファ1.2のRAM領域はQバッファ1.0における入力文字列の符号化が1つ加わる毎に符号化済み文字数分の文字がRAM領域の左端から捨てられ、Qバッファ1.0の符号化済み文字列がRAM領域の右端よりシフトして格納される。

【0061】図8は図2のデータ圧縮部3.0による符号化処理を示したフローチャートであり、次のようにして行われる。

S1:図6のステップS1と同様、Pバッファ1.2の前半に予め作成された初期値文字列(n文字)を格納する。

S2:符号語を入力し、複製モードのときはPバッファ1.2を参照して文字列を復元する。

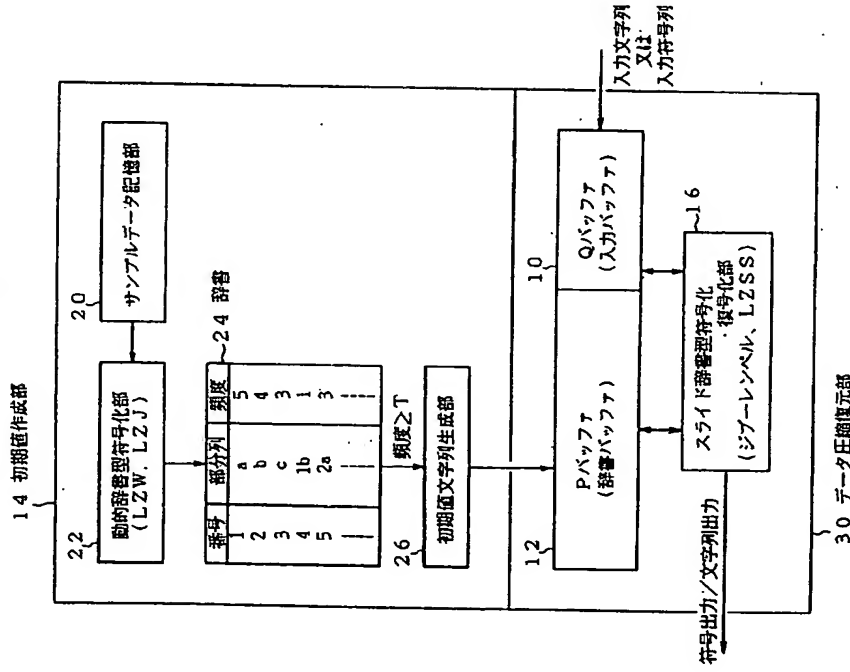
【0062】S3:図6のステップS3と同様に、Pバッファ1.2の削除と登録の処理を行う。以下同様にして、ステップS2、S3の処理を繰り返して、符号化した文字列を復元する。図9は本発明の初期値文字列の作成処理の他の実施例を示した説明図である。図9の初期値文字列の作成処理にあっては、まずスライド辞書型アルゴリズムに従ってサンプルデータを符号化し、サンプルデータから得られた符号列を対象に動的辞書型アルゴリズムに従った符号化を行い、この符号化の際に作成される辞書の木構造における節点にカウントを設けて使用頻度を計数し、符号化が終了したならば所定の閾値T以上の使

用頻度をもつ文字列を取り出し、1列に並べて初期値文字列を作成する。

【0063】即ち、まずサンプルデータをQバッファ1.0aに入力し、Pバッファ1.2hの登録済み文字列を検索して最長一致する文字列を求め、開始位置p1と一致長q1で符号化する。勿論、この符号化は2文字以上の場合に行われ、1文字の場合は生データを出力する。このようにスライド辞書型アルゴリズムによる符号化で出力される符号列S1、S2、...、Si、...、Sj、...、Sj+1、...、Sj+2、...、Sj+3、...、Sj+4、...、Sj+5、...、Sj+6、...、Sj+7、...、Sj+8、...、Sj+9、...、Sj+10、...、Sj+11、...、Sj+12、...、Sj+13、...、Sj+14、...、Sj+15、...、Sj+16、...、Sj+17、...、Sj+18、...、Sj+19、...、Sj+20、...、Sj+21、...、Sj+22、...、Sj+23、...、Sj+24、...、Sj+25、...、Sj+26、...、Sj+27、...、Sj+28、...、Sj+29、...、Sj+30、...、Sj+31、...、Sj+32、...、Sj+33、...、Sj+34、...、Sj+35、...、Sj+36、...、Sj+37、...、Sj+38、...、Sj+39、...、Sj+40、...、Sj+41、...、Sj+42、...、Sj+43、...、Sj+44、...、Sj+45、...、Sj+46、...、Sj+47、...、Sj+48、...、Sj+49、...、Sj+50、...、Sj+51、...、Sj+52、...、Sj+53、...、Sj+54、...、Sj+55、...、Sj+56、...、Sj+57、...、Sj+58、...、Sj+59、...、Sj+60、...、Sj+61、...、Sj+62、...、Sj+63、...、Sj+64、...、Sj+65、...、Sj+66、...、Sj+67、...、Sj+68、...、Sj+69、...、Sj+70、...、Sj+71、...、Sj+72、...、Sj+73、...、Sj+74、...、Sj+75、...、Sj+76、...、Sj+77、...、Sj+78、...、Sj+79、...、Sj+80、...、Sj+81、...、Sj+82、...、Sj+83、...、Sj+84、...、Sj+85、...、Sj+86、...、Sj+87、...、Sj+88、...、Sj+89、...、Sj+90、...、Sj+91、...、Sj+92、...、Sj+93、...、Sj+94、...、Sj+95、...、Sj+96、...、Sj+97、...、Sj+98、...、Sj+99、...、Sj+100、...、Sj+101、...、Sj+102、...、Sj+103、...、Sj+104、...、Sj+105、...、Sj+106、...、Sj+107、...、Sj+108、...、Sj+109、...、Sj+110、...、Sj+111、...、Sj+112、...、Sj+113、...、Sj+114、...、Sj+115、...、Sj+116、...、Sj+117、...、Sj+118、...、Sj+119、...、Sj+120、...、Sj+121、...、Sj+122、...、Sj+123、...、Sj+124、...、Sj+125、...、Sj+126、...、Sj+127、...、Sj+128、...、Sj+129、...、Sj+130、...、Sj+131、...、Sj+132、...、Sj+133、...、Sj+134、...、Sj+135、...、Sj+136、...、Sj+137、...、Sj+138、...、Sj+139、...、Sj+140、...、Sj+141、...、Sj+142、...、Sj+143、...、Sj+144、...、Sj+145、...、Sj+146、...、Sj+147、...、Sj+148、...、Sj+149、...、Sj+150、...、Sj+151、...、Sj+152、...、Sj+153、...、Sj+154、...、Sj+155、...、Sj+156、...、Sj+157、...、Sj+158、...、Sj+159、...、Sj+160、...、Sj+161、...、Sj+162、...、Sj+163、...、Sj+164、...、Sj+165、...、Sj+166、...、Sj+167、...、Sj+168、...、Sj+169、...、Sj+170、...、Sj+171、...、Sj+172、...、Sj+173、...、Sj+174、...、Sj+175、...、Sj+176、...、Sj+177、...、Sj+178、...、Sj+179、...、Sj+180、...、Sj+181、...、Sj+182、...、Sj+183、...、Sj+184、...、Sj+185、...、Sj+186、...、Sj+187、...、Sj+188、...、Sj+189、...、Sj+190、...、Sj+191、...、Sj+192、...、Sj+193、...、Sj+194、...、Sj+195、...、Sj+196、...、Sj+197、...、Sj+198、...、Sj+199、...、Sj+200、...、Sj+201、...、Sj+202、...、Sj+203、...、Sj+204、...、Sj+205、...、Sj+206、...、Sj+207、...、Sj+208、...、Sj+209、...、Sj+210、...、Sj+211、...、Sj+212、...、Sj+213、...、Sj+214、...、Sj+215、...、Sj+216、...、Sj+217、...、Sj+218、...、Sj+219、...、Sj+220、...、Sj+221、...、Sj+222、...、Sj+223、...、Sj+224、...、Sj+225、...、Sj+226、...、Sj+227、...、Sj+228、...、Sj+229、...、Sj+230、...、Sj+231、...、Sj+232、...、Sj+233、...、Sj+234、...、Sj+235、...、Sj+236、...、Sj+237、...、Sj+238、...、Sj+239、...、Sj+240、...、Sj+241、...、Sj+242、...、Sj+243、...、Sj+244、...、Sj+245、...、Sj+246、...、Sj+247、...、Sj+248、...、Sj+249、...、Sj+250、...、Sj+251、...、Sj+252、...、Sj+253、...、Sj+254、...、Sj+255、...、Sj+256、...、Sj+257、...、Sj+258、...、Sj+259、...、Sj+260、...、Sj+261、...、Sj+262、...、Sj+263、...、Sj+264、...、Sj+265、...、Sj+266、...、Sj+267、...、Sj+268、...、Sj+269、...、Sj+270、...、Sj+271、...、Sj+272、...、Sj+273、...、Sj+274、...、Sj+275、...、Sj+276、...、Sj+277、...、Sj+278、...、Sj+279、...、Sj+280、...、Sj+281、...、Sj+282、...、Sj+283、...、Sj+284、...、Sj+285、...、Sj+286、...、Sj+287、...、Sj+288、...、Sj+289、...、Sj+290、...、Sj+291、...、Sj+292、...、Sj+293、...、Sj+294、...、Sj+295、...、Sj+296、...、Sj+297、...、Sj+298、...、Sj+299、...、Sj+300、...、Sj+301、...、Sj+302、...、Sj+303、...、Sj+304、...、Sj+305、...、Sj+306、...、Sj+307、...、Sj+308、...、Sj+309、...、Sj+310、...、Sj+311、...、Sj+312、...、Sj+313、...、Sj+314、...、Sj+315、...、Sj+316、...、Sj+317、...、Sj+318、...、Sj+319、...、Sj+320、...、Sj+321、...、Sj+322、...、Sj+323、...、Sj+324、...、Sj+325、...、Sj+326、...、Sj+327、...、Sj+328、...、Sj+329、...、Sj+330、...、Sj+331、...、Sj+332、...、Sj+333、...、Sj+334、...、Sj+335、...、Sj+336、...、Sj+337、...、Sj+338、...、Sj+339、...、Sj+340、...、Sj+341、...、Sj+342、...、Sj+343、...、Sj+344、...、Sj+345、...、Sj+346、...、Sj+347、...、Sj+348、...、Sj+349、...、Sj+350、...、Sj+351、...、Sj+352、...、Sj+353、...、Sj+354、...、Sj+355、...、Sj+356、...、Sj+357、...、Sj+358、...、Sj+359、...、Sj+360、...、Sj+361、...、Sj+362、...、Sj+363、...、Sj+364、...、Sj+365、...、Sj+366、...、Sj+367、...、Sj+368、...、Sj+369、...、Sj+370、...、Sj+371、...、Sj+372、...、Sj+373、...、Sj+374、...、Sj+375、...、Sj+376、...、Sj+377、...、Sj+378、...、Sj+379、...、Sj+380、...、Sj+381、...、Sj+382、...、Sj+383、...、Sj+384、...、Sj+385、...、Sj+386、...、Sj+387、...、Sj+388、...、Sj+389、...、Sj+390、...、Sj+391、...、Sj+392、...、Sj+393、...、Sj+394、...、Sj+395、...、Sj+396、...、Sj+397、...、Sj+398、...、Sj+399、...、Sj+400、...、Sj+401、...、Sj+402、...、Sj+403、...、Sj+404、...、Sj+405、...、Sj+406、...、Sj+407、...、Sj+408、...、Sj+409、...、Sj+410、...、Sj+411、...、Sj+412、...、Sj+413、...、Sj+414、...、Sj+415、...、Sj+416、...、Sj+417、...、Sj+418、...、Sj+419、...、Sj+420、...、Sj+421、...、Sj+422、...、Sj+423、...、Sj+424、...、Sj+425、...、Sj+426、...、Sj+427、...、Sj+428、...、Sj+429、...、Sj+430、...、Sj+431、...、Sj+432、...、Sj+433、...、Sj+434、...、Sj+435、...、Sj+436、...、Sj+437、...、Sj+438、...、Sj+439、...、Sj+440、...、Sj+441、...、Sj+442、...、Sj+443、...、Sj+444、...、Sj+445、...、Sj+446、...、Sj+447、...、Sj+448、...、Sj+449、...、Sj+450、...、Sj+451、...、Sj+452、...、Sj+453、...、Sj+454、...、Sj+455、...、Sj+456、...、Sj+457、...、Sj+458、...、Sj+459、...、Sj+460、...、Sj+461、...、Sj+462、...、Sj+463、...、Sj+464、...、Sj+465、...、Sj+466、...、Sj+467、...、Sj+468、...、Sj+469、...、Sj+470、...、Sj+471、...、Sj+472、...、Sj+473、...、Sj+474、...、Sj+475、...、Sj+476、...、Sj+477、...、Sj+478、...、Sj+479、...、Sj+480、...、Sj+481、...、Sj+482、...、Sj+483、...、Sj+484、...、Sj+485、...、Sj+486、...、Sj+487、...、Sj+488、...、Sj+489、...、Sj+490、...、Sj+491、...、Sj+492、...、Sj+493、...、Sj+494、...、Sj+495、...、Sj+496、...、Sj+497、...、Sj+498、...、Sj+499、...、Sj+500、...、Sj+501、...、Sj+502、...、Sj+503、...、Sj+504、...、Sj+505、...、Sj+506、...、Sj+507、...、Sj+508、...、Sj+509、...、Sj+510、...、Sj+511、...、Sj+512、...、Sj+513、...、Sj+514、...、Sj+515、...、Sj+516、...、Sj+517、...、Sj+518、...、Sj+519、...、Sj+520、...、Sj+521、...、Sj+522、...、Sj+523、...、Sj+524、...、Sj+525、...、Sj+526、...、Sj+527、...、Sj+528、...、Sj+529、...、Sj+530、...、Sj+531、...、Sj+532、...、Sj+533、...、Sj+534、...、Sj+535、...、Sj+536、...、Sj+537、...、Sj+538、...、Sj+539、...、Sj+540、...、Sj+541、...、Sj+542、...、Sj+543、...、Sj+544、...、Sj+545、...、Sj+546、...、Sj+547、...、Sj+548、...、Sj+549、...、Sj+550、...、Sj+551、...、Sj+552、...、Sj+553、...、Sj+554、...、Sj+555、...、Sj+556、...、Sj+557、...、Sj+558、...、Sj+559、...、Sj+560、...、Sj+561、...、Sj+562、...、Sj+563、...、Sj+564、...、Sj+565、...、Sj+566、...、Sj+567、...、Sj+568、...、Sj+569、...、Sj+570、...、Sj+571、...、Sj+572、...、Sj+573、...、Sj+574、...、Sj+575、...、Sj+576、...、Sj+577、...、Sj+578、...、Sj+579、...、Sj+580、...、Sj+581、...、Sj+582、...、Sj+583、...、Sj+584、...、Sj+585、...、Sj+586、...、Sj+587、...、Sj+588、...、Sj+589、...、Sj+590、...、Sj+591、...、Sj+592、...、Sj+593、...、Sj+594、...、Sj+595、...、Sj+596、...、Sj+597、...、Sj+598、...、Sj+599、...、Sj+600、...、Sj+601、...、Sj+602、...、Sj+603、...、Sj+604、...、Sj+605、...、Sj+606、...、Sj+607、...、Sj+608、...、Sj+609、...、Sj+610、...、Sj+611、...、Sj+612、...、Sj+613、...、Sj+614、...、Sj+615、...、Sj+616、...、Sj+617、...、Sj+618、...、Sj+619、...、Sj+620、...、Sj+621、...、Sj+622、...、Sj+623、...、Sj+624、...、Sj+625、...、Sj+626、...、Sj+627、...、Sj+628、...、Sj+629、...、Sj+630、...、Sj+631、...、Sj+632、...、Sj+633、...、Sj+634、...、Sj+635、...、Sj+636、...、Sj+637、...、Sj+638、...、Sj+639、...、Sj+640、...、Sj+641、...、Sj+642、...、Sj+643、...、Sj+644、...、Sj+645、...、Sj+646、...、Sj+647、...、Sj+648、...、Sj+649、...、Sj+650、...、Sj+651、...、Sj+652、...、Sj+653、...、Sj+654、...、Sj+655、...、Sj+656、...、Sj+657、...、Sj+658、...、Sj+659、...、Sj+660、...、Sj+661、...、Sj+662、...、Sj+663、...、Sj+664、...、Sj+665、...、Sj+666、...、Sj+667、...、Sj+668、...、Sj+669、...、Sj+670、...、Sj+671、...、Sj+672、...、Sj+673、...、Sj+674、...、Sj+675、...、Sj+676、...、Sj+677、...、Sj+678、...、Sj+679、...、Sj+680、...、Sj+681、...、Sj+682、...、Sj+683、...、Sj+684、...、Sj+685、...、Sj+686、...、Sj+687、...、Sj+688、...、Sj+689、...、Sj+690、...、Sj+691、...、Sj+692、...、Sj+693、...、Sj+694、...、Sj+695、...、Sj+696、...、Sj+697、...、Sj+698、...、Sj+699、...、Sj+700、...、Sj+701、...、Sj+702、...、Sj+703、...、Sj+704、...、Sj+705、...、Sj+706、...、Sj+707、...、Sj+708、...、Sj+709、...、Sj+710、...、Sj+711、...、Sj+712、...、Sj+713、...、Sj+714、...、Sj+715、...、Sj+716、...、Sj+717、...、Sj+718、...、Sj+719、...、Sj+720、...、Sj+721、...、Sj+722、...、Sj+723、...、Sj+724、...、Sj+725、...、Sj+726、...、Sj+727、...、Sj+728、...、Sj+729、...、Sj+730、...、Sj+731、...、Sj+732、...、Sj+733、...、Sj+734、...、Sj+735、...、Sj+736、...、Sj+737、...、Sj+738、...、Sj+739、...、Sj+740、...、Sj+741、...、Sj+742、...、Sj+743、...、Sj+744、...、Sj+745、...、Sj+746、...、Sj+747、...、Sj+748、...、Sj+749、...、Sj+750、...、Sj+751、...、Sj+752、...、Sj+753、...、Sj+754、...、Sj+755、...、Sj+756、...、Sj+757、...、Sj+758、...、Sj+759、...、Sj+760、...、Sj+761、...、Sj+762、...、Sj+763、...、Sj+764、...、Sj+765、...、Sj+766、...、Sj+767、...、Sj+768、...、Sj+769、...、Sj+770、...、Sj+771、...、Sj+772、...、Sj+773、...、Sj+774、...、Sj+775、...、Sj+776、...、Sj+777、...、Sj+778、...、Sj+779、...、Sj+780、...、Sj+781、...、Sj+782、...、Sj+783、...、Sj+784、...、Sj+785、...、Sj+786、...、Sj+787、...、Sj+788、...、Sj+789、...、Sj+790、...、Sj+791、...、Sj+792、...、Sj+793、...、Sj+794、...、Sj+795、...、Sj+796、...、Sj+797、...、Sj+798、...、Sj+799、...、Sj+800、...、Sj+801、...、Sj+802、...、Sj+803、...、Sj+804、...、Sj+805、...、Sj+806、...、Sj+807、...、Sj+808、...、Sj+809、...、Sj+810、...、Sj+811、...、Sj+812、...、Sj+813、...、Sj+814、...、Sj+815、...、Sj+816、...、Sj+817、...、Sj+818、...、Sj+819、...、Sj+820、...、Sj+821、...、Sj+822、...、Sj+823、...、Sj+824、...、Sj+825、...、Sj+826、...、Sj+827、...、Sj+828、...、Sj+829、...、Sj+830、...、Sj+831、...、Sj+832、...、Sj+833、...、Sj+834、...、Sj+835、...、Sj+836、...、Sj+837、...、Sj+838、...、Sj+839、...、Sj+840、...、Sj+841、...、Sj+842、...、Sj+843、...、Sj+844、...、Sj+845、...、Sj+846、...、Sj+847、...、Sj+848、...、Sj+849、...、Sj+850、...、Sj+851、...、Sj+852、...、Sj+853、...、Sj+854、...、Sj+855、...、Sj+856、...、Sj+857、...、Sj+858、...、Sj+859、...、Sj+860、...、Sj+861、...、Sj+862、...、Sj+863、...、Sj+864、...、Sj+865、...、Sj+866、...、Sj+867、...、Sj+868、...、Sj+869、...、Sj+870、...、Sj+871、...、Sj+872、...、Sj+873、...、Sj+874、...、Sj+875、...、Sj+876、...、Sj+877、...、Sj+878、...、Sj+879、...、Sj+880、...、Sj+881、...、Sj+882、...、Sj+883、...、Sj+884、...、Sj+885、...、Sj+886、...、Sj+887、...、Sj+888、...、Sj+889、...、Sj+890、...、Sj+891、...、Sj+892、...、Sj+893、...、Sj+894、...、Sj+895、...、Sj+896、...、Sj+897、...、Sj+898、...、Sj+899、...、Sj+900、...、Sj+901、...、Sj+902、...、Sj+903、...、Sj+904、...、Sj+905、...、Sj+906、...、Sj+907、...、Sj+908、...、Sj+909、...、Sj+910、...、Sj+911、...、Sj+912、...、Sj+913、...、Sj+914、...、Sj+915、...、Sj+916、...、Sj+917、...、Sj+918、...、Sj+919、...、Sj+920、...、Sj+921、...、Sj+922、...、Sj+923、...、Sj+924、...、Sj+925、...、Sj+926、...、Sj+927、...、Sj+928、...、Sj+929、...、Sj+930、...、Sj+931、...、Sj+932、...、Sj+933、...、Sj+934、...、Sj+935、...、Sj+936、...、Sj+937、...、Sj+938、...、Sj+939、...、Sj+940、...、Sj+941、...、Sj+942、...、Sj+943、...、Sj+944、...、Sj+945、...、Sj+946、...、Sj+947、...、Sj+948、...、Sj+949、...、Sj+950、...、Sj+951、...、Sj+952、...、Sj+953、...、Sj+954、...、Sj+955、...、Sj+956、...、Sj+957、...、Sj+958、...、Sj+959、...、Sj+960、...、Sj+961、...、Sj+962、...、Sj+963、...、Sj+964、...、Sj+965、...、Sj+966、...、Sj+967、...、Sj+968、...、Sj+969、...、Sj+970、...、Sj+971、...、Sj+972、...、Sj+973、...、Sj+974、...、Sj+975、...、Sj+976、...、Sj+977、...、Sj+978、...、Sj+979、...、Sj+980、...、Sj+981、...、Sj+982、...、Sj+983、...、Sj+984、...、Sj+985、...、Sj+986、...、Sj+987、...、Sj+988、...、Sj+989、...、Sj+990、...、Sj+991、...、Sj+992、...、Sj+993、...、Sj+994、...、Sj+995、...、Sj+996、...、Sj+997、...、Sj+998、...、Sj+999、...、Sj+1000、...、Sj+1001、...、Sj+1002、...、Sj+1003、...、Sj+1004、...、Sj+1005、...、Sj+1006、...、Sj+1007、...、Sj+1008、...、Sj+

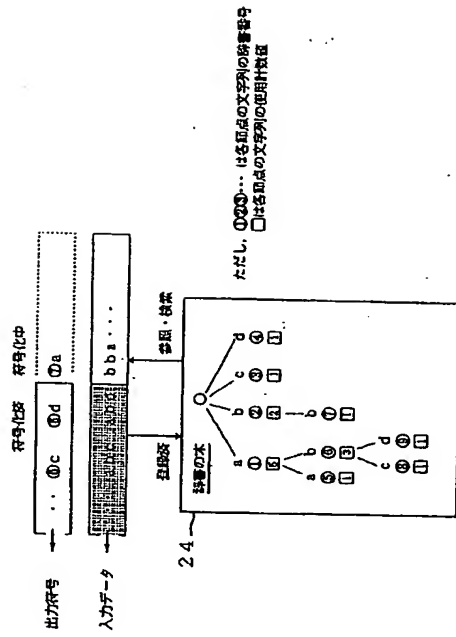
【図2】

本発明の実施例構成図



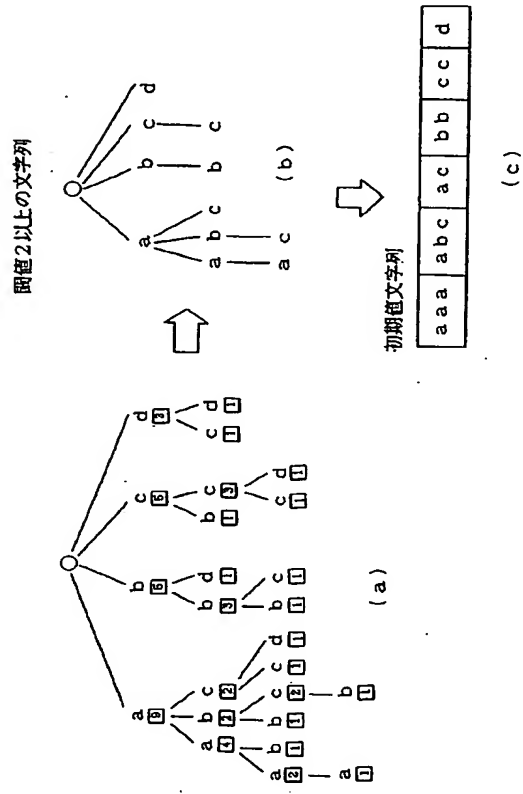
【図4】

本発明のサンプルデータを対象とした動的辞書型符号化と辞書の文字列使用回数
の計数を示した説明図



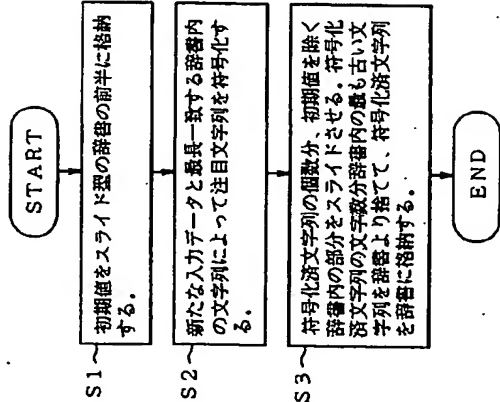
【図5】

本発明の符号化辞書の使用頻度に基づく初期値文字列の生成を示した説明図



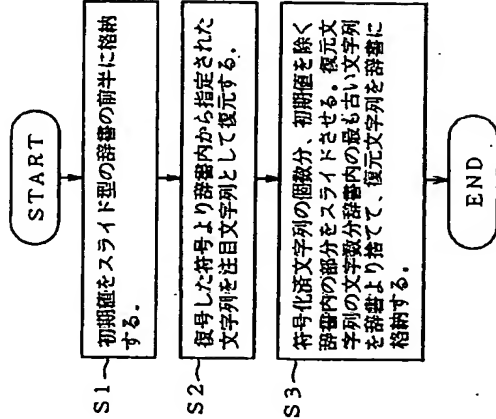
【図6】

本発明の初期値文字列を用いたスライド辞書型符号化のフローチャート



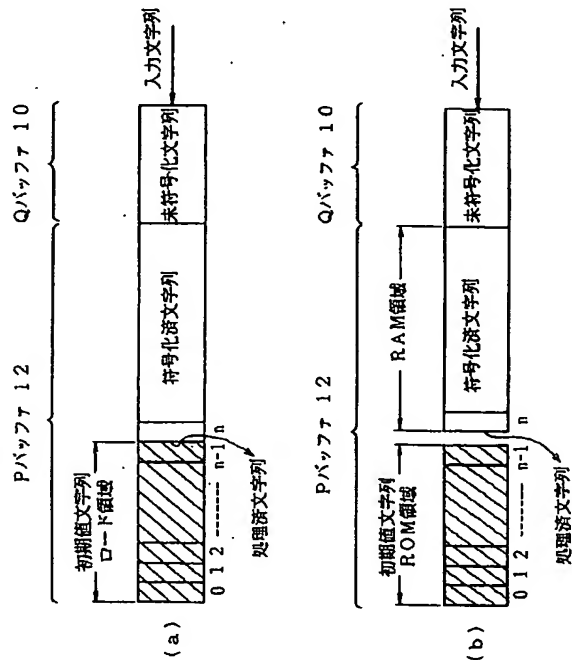
【図8】

本発明のスライド辞書型復号化のフローチャート



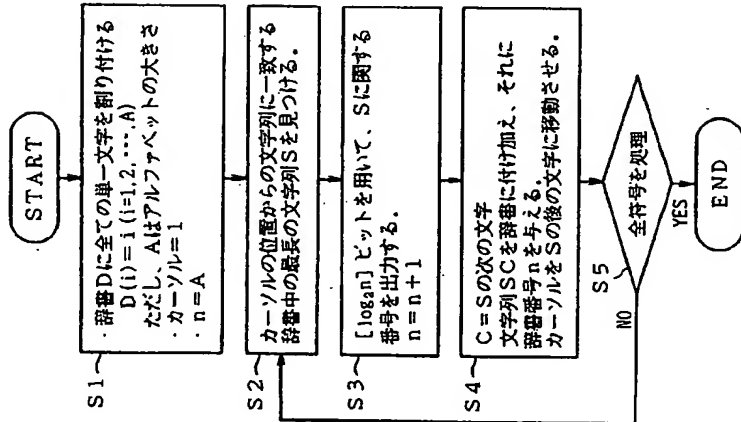
【図7】

本発明のスライド辞書型符号化で使用するPバッファの構成を示した説明図



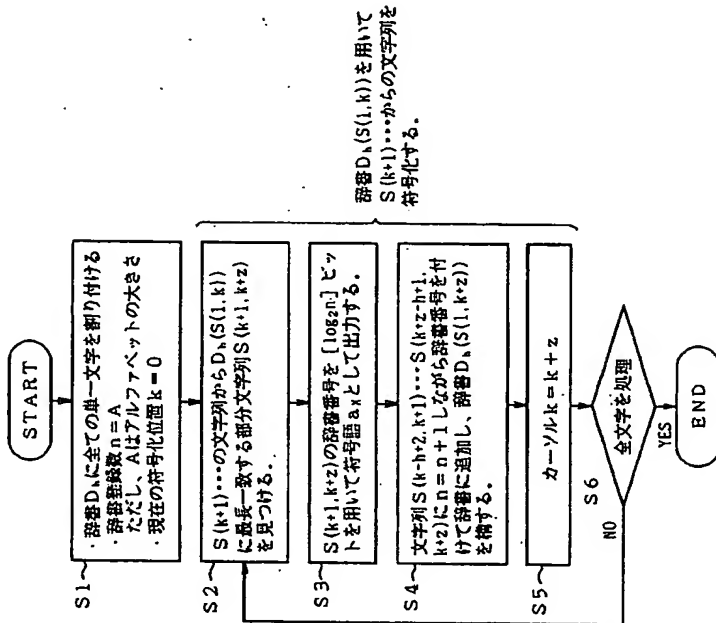
【図15】

従来のLZW復号化アルゴリズムを示したフローチャート



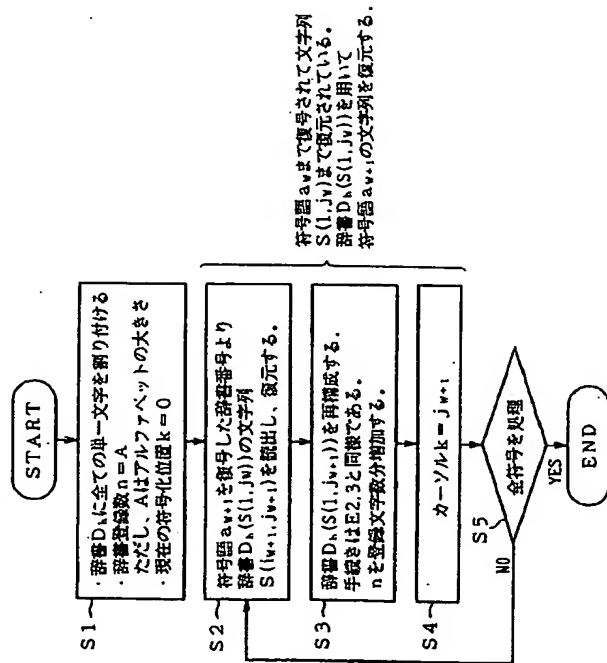
【図16】

従来のLZJ符号化アルゴリズムを示したフローチャート

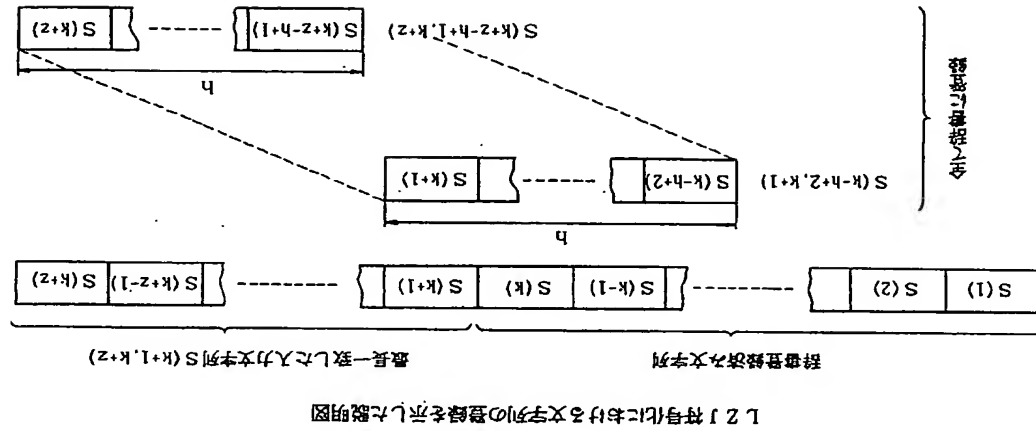


【図17】

従来のL2J復号化アルゴリズムを示したフローチャート



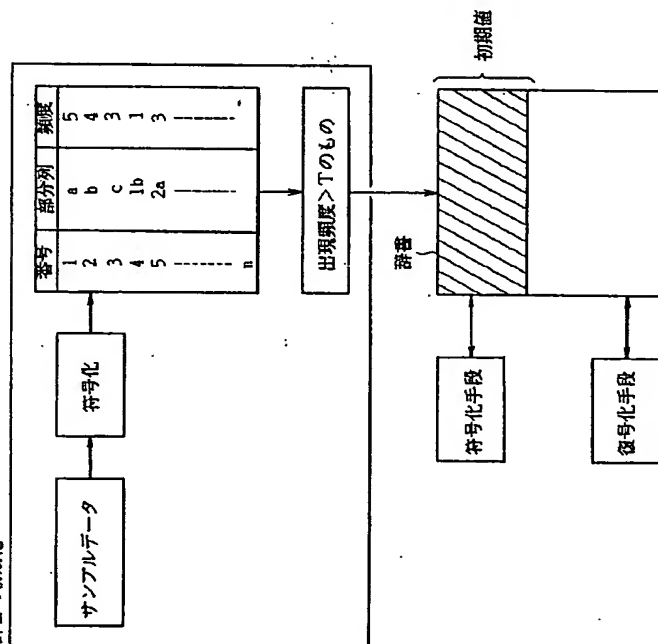
【図18】



【図19】

本発明者が既に提案しているLZW符号を用いたデータ圧縮における辞書の初期登録の説明図

辞書の初期化



フロントページの続き

(72)発明者 千葉 広隆
 神奈川県川崎市中原区上小田中1015番地
 富士通株式会社内